Conference Paper

A model for coupling multi-agent social interactions and traffic simulation

Author(s):
Hackney, Jeremy K.; Marchal, Fabrice

Publication Date:
2008-07

Permanent Link:
https://doi.org/10.3929/ethz-a-005652382

Rights / License:
In Copyright - Non-Commercial Use Permitted
A Model for Coupling Multi-Agent Social Interactions and Traffic Simulation

Submission for Presentation to the Transportation Research Board 2009
A Model for Coupling Multi-Agent Social Interactions and Traffic Simulation

Jeremy Hackney
Institute for Transport Planning and Systems/ETH
Zurich

Fabrice Marchal
CNRS Associate Researcher
Lab of Transportation Economics (LET)
Lyon

Phone: +41(0)446333325
Fax: +41(0)446331057
email: hackney@ivt.baug.ethz.ch

July 2008

Abstract

Understanding social and travel interdependencies becomes more important as schedules and work locations become more flexible. In lieu of comprehensive empirical studies of these interactions, agent simulation is available to construct hypothetical interdependencies between travellers and activity-travel choices. A general geographic social interaction model, based on the Multi-Agent Transportation Simulation Toolbox (MatSim-T), is introduced to serve as a laboratory for scenario- and hypothesis testing. The basic tools of transportation science: utility maximization, activity plans, and generalized travel cost, are used to construct social networks for a geographically distributed population of tens of thousands of agents. The social networks are used in static or dynamic configurations in optimization iterations to modify travel demand (location, activity, and/or timing). This paper summarizes the model and presents detailed validation results, describing the individual effects of socializing assumptions on travel behavior.

Keywords

behavior microsimulation – social networks – activity-based planning – travel behavior
1. The interactive traveller and transportation behavior

The fact that people plan and undertake trips and activities jointly, depending on the trips and activities of other people, is recognized in the transportation demand field but is usually disregarded for several reasons (Kurani and Kitamura 1996). Supply design criteria necessitate satisfying the highest demand hours, which typically occur under conditions of the work commute, a trip traditionally well explained by models of individuals. The convenient econometric model estimation tools assume independent observations. Finally, there is a lack of datasets detailing the social dynamics and their significance (Manski 2000).

Considering behavior to be entirely independent succeeds for regularly occurring, inelastic behavior, like work (non-discretionary) trips. Travel preferences are explained by sociodemographics, generalized travel costs and destination attributes; other processes take either a long-term character exogenous to the model (Hensher 2002), or they play an insignificant role due to the limited flexibility of the traveller. Discretionary behaviors are commonly explained by adding regression variables like taste heterogeneity and cohort or habitual behavior effects. While improving the statistical specification and the conceptual validity of the models, this neglects accounting for traveller interactions which are known to be real influences. This rules the models out for studying processes of real behavior, and may limit their usefulness in prediction, as well. Taking a lesson from the field of network economics, the model parameters and output without interactions may be plausible, but this may be for the wrong reason (Shy 2001).

Endogenizing interactions explicitly accounts for who influences whom, which adjusts the supply and demand equilibrium. But the drawback to methods using social interactions is that it becomes necessary to know the specific social connections (social network), as well as the direction and strength of influence on an individual's decision making that is communicated by this network (Manski 1993; Bramoullé et al. 2007).

There is a field of interest expanding in geographical social networks and their integration with transportation and communications over distance. Carrasco and Miller (2006), Silvis, et al. (2006), Axhausen and Frei (2007), Mok, et al. (2007), Carrasco et al. (2008), and Carrasco et al. (Forthcoming) present new empirical research focused on travel and communications. Christakis and Fowler (2007) and (2008), and Rothenberg, et al. (2005) have mapped and modelled social networks and spatial interactions that correlate with health-related behaviors. Arentze and Timmermans (2006) implement a comprehensive model for ego-centric social interactions and activity patterns. Small-scale toy simulations of travel and socializing are presented in Hackney and Axhausen (2006) and Marchal and Nagel (2006). Estimations of
socially-dependent mode/destination travel models have been investigated for example by Páez and Scott (2004), Dugundji and Gulyas (2003), and Dugundji and Walker (2005).

1.1 The contribution of this work

Because the sparse observations of social interactions in the geographical or transportation context do not permit deep study of the interplay between geography and socializing, this work presents a scenario-based approach to generate datasets based on plausible interdependencies between social and travel behavior. A microsimulation of travelling agent actors is used to postulate different social interactions and observe the feedbacks on short-term travel decisions.

Experiments combine algorithms and/or datasets for the three main elements of study: socializing, geography, and travel behavior, with adjustable system coupling. The result is a set of geographically-embedded social networks to study, and a set of travel behavior patterns that depend on social interactions that have been well-defined.

Using multi-agent simulations to model inter-actor relationships that lead to macro-scale system properties is a combination of deductive and inductive methods, sometimes called “generative science” (Sawyer 2004). Agent modeling allows researchers to control and experiment with microscopic behavior and observe the emergent macroscopic system (Bankes 1993; Axtell 2000). The acceptance of the results depends on clear explanation of the assumptions and thorough model validation.

This paper describes the implementation of the social network module within travel microsimulation framework and shows a battery of validation results, concluding with plans for extensive probing of the models, further refinement, and possible applications.
2. Social Networks Module and MATSim

This work builds on the simulations of joint travel behavior of Hackney and Axhausen (2006) and Marchal and Nagel (2006) by adding flexibility to experiment with large social networks embedded in various systems of transportation networks and time-space geographies of activities. The social network module extends the Java-implemented Multi-Agent Transportation Simulation Toolbox (MATSim, Rieser et al. 2007)). The toolbox offers the advantage of efficient data structures, modularity and object-oriented structure, management of the I/O and run history, and parallel-processed dynamic traffic assignment.

MATSim is initialized with a geographic World, a transportation Network, a set of Facilities in which activities are carried out, and a set of agents with Plans to be carried out (Balmer 2007). The plans consist of each agent's activities and their locations, generated using detailed datasets of travel behavior. A loop iteratively seeks maximum utility for each agent's day plan by an evolutionary algorithm which favors successful outcomes of the agents' adaptation strategies (Figure 1). Agents gain utility for participating in activities and disutility for travel and delay. In a "re-planning" phase the agents can alter either their route, their departure time, or nothing at all. Altering the order, type, and number of activities or their locations are currently not standard replanning strategies, but a location choice module is introduced here in the framework of social network exchanges. The stopping criterion is the stability of the utility score, indicating Nash equilibrium. MATSim's focus is on a single day's travel, though other short-term horizons would not be infeasible. The iterative process of utility maximization should be viewed as an optimization algorithm seeking system relaxation given a certain utility function, rather than a learning process or a sequence of time steps.
A personal social network, or "EgoNetwork" object keeps track of acquaintances, or "friends", extending the mental map of the MATSim agent, which stores activity locations. Here, a "friend" is a general reference to adjacency in a matrix of agents. The researcher can access social links and agents from the EgoNetwork using either piece of information. The links have attributes Euclidean length, strength, to/from vertex (= agents), and the iteration and activity type where they were established and when they were activated last (their most recent face to face encounter).

A "SocialNetwork" object consists of a list of pointers to the EgoNetworks, thus constructing a global social network. The SocialNetwork class defines whether links are directed or undirected, and contains methods to initiate, strengthen, and remove links, and a saturation function to simulate the cognitive limitations that prevent humans from maintaining too many friends. Various algorithms from the literature are included to initialize the social network: small world (Watts 1999; Newman et al. 2002), scale-free (Barabasi et al. 1999), random (e.g. Dorogovtsev and Mendes 2003), as well as spatially embedded versions of these (Penrose 1991; Andrade et al. 2005; Wong et al. 2006).

The day plans of the agents can be affected by socializing in two ways. First, agents can exchange information to modify their mental maps and thus knowledge of choice sets. In the examples presented here, agents exchange knowledge about secondary locations and then draw on this knowledge to attempt to improve their plans in the optimization iterations. Information exchange may be defined to occur in space (time- and space-dependent encounters like communication of disease, personal transactions, working together) or nonspatially (models to encompass all other exchanges of knowledge which do require
explicit travel in this day's plan: general shared knowledge, commonly held friends, etc.) (Hackney 2007).

Second, the utility function can be modified to include benefits or penalties for different types of social behavior as described in the economics and sociology literature. An example is shown where the utility of leisure activities increases with the log of the number of friends present. It is important to keep in mind that the agent model is optimized with respect to the parameters included in the utility function and thus only represents the tradeoffs between these parameters: socializing versus travel time, for example. Nothing else in the system is explicitly constrained, and these behaviors thus "emerge" consistent with the system configuration as the agents maximize their utility.

Finally, the package includes a statistics module using the JUNG network statistics package (O'Madadhain et al. 2005) that outputs social, geographic, and travel statistics for each agent, each edge, and for the graph as a whole, for each iteration step. Graphical output of ego networks and activity spaces can be written to KML format and the social networks can be output to Pajek (Batagelj and Mrvar 2003).
3. Sample application

Social networks can be used in three kinds of experiments: first, the researcher can choose a fixed set of agent plans and allow agents who meet face to face to establish relationships with one another, perhaps augmented by other non-spatial social processes. This generates a geographically embedded social network based on face-to-face meetings and some algorithm for befriending under that circumstance. Second, a social network can be generated using any number of algorithms and then frozen, to be used as a static influence in optimizing the activity plans. The goal here is to try different interactions and to compare the resulting activity plans with those that do not use interdependent agents. Third, the social network and the plans can be coupled and altered together according to the researcher's hypotheses about the interdependence of activities and social connections. This does not represent an evolution of social networks in time, but a relaxation of the social-geographic system with respect to the utility function.

Seven different social network interaction models are described and their social network and travel behavior results are compared to illustrate how the module functions and what effect experiments have on the travel behavior simulation.

3.1 Model description

All model configurations have the following in common:

1) 1% sample of the population in the Zurich region (car mode only) = 8760 agents and 6438 geocoded facilities with opening times (Meister et al. 2008)

2) 5 activity types: home, work, education, shopping, leisure, with activity chains drawn from micro-census travel data (Swiss Federal Statistical Office 2001). The population makes 3.38 trips per day.

3) 100km x 100km region surrounding Zurich (agents' initial plans were chosen such that the entire day plan takes place within this region).

4) Swiss National Road Network modified to 1% of capacity (Vrtic et al. 2003)
<table>
<thead>
<tr>
<th>Name</th>
<th>Type</th>
<th>Social network</th>
<th>Social interact</th>
<th>Score</th>
<th>Replan</th>
<th>Iterations</th>
<th>Runtime</th>
</tr>
</thead>
<tbody>
<tr>
<td>Config1</td>
<td>reference</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>500</td>
<td>1.7GB, 11.5hrs</td>
</tr>
<tr>
<td>Config7</td>
<td>reference</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>500</td>
<td>1.7GB, 15hrs</td>
</tr>
<tr>
<td>Config2</td>
<td>SN_time</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>500</td>
<td>3.5GB, 14.5hrs</td>
</tr>
<tr>
<td>Config3</td>
<td>SN_loc</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>500</td>
<td>3.9-4.2GB, 21hours</td>
</tr>
<tr>
<td>Config4</td>
<td>SN_loc</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>500</td>
<td>3.9-4.2GB, 21hours</td>
</tr>
<tr>
<td>Config5</td>
<td>SN_loc_Dyn_1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>500</td>
<td>3.9-4.2GB, 21hours</td>
</tr>
<tr>
<td>Config6</td>
<td>SN_loc_Dyn_2</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>500</td>
<td>3.9-4.2GB, 21hours</td>
</tr>
</tbody>
</table>

Social network:
0 = none;
1 = constant undirected Erdős/Rényi graph of average degree 12 with link probability attenuating with distance $^{-1.5}$;
2 = initialization as in 1) with 5% chance of befriending a randomly encountered agent and 5% chance of random link removal to maintain average degree = 12.

Social interactions:
0 = none;
1 = one facility location is exchanged between befriended agents per iteration

Scoring:
0 = standard utility function;
1 = standard utility + constant * \( \ln(\text{number of friends present}) \) * (1,0) leisure dummy

Replan:
0 = [15% activity time mutation, 15% re-route, 70% logistic based on score];
1 = [10% activity time mutation, 10% re-route, 10% re-locate shopping or leisure activity from agent knowledge, 70% logistic based on score];
2 = [10% activity time mutation, 10% re-route, 10% re-locate shopping or leisure activity from pool of all facilities, 70% logistic based on score]

Activity location choice involves switching one shopping or leisure location per iteration.

Runtime: 2.2 GHz processor, single thread
Config1 is a base case corresponding to a standard MATSim configuration with independently utility-maximizing agents. There is no social network and the agents can only alter routes and departure times to improve their utility. Agents adjust their day plans using one of the listed strategies allocated to them randomly according to the percentage given in the table. Utility is determined by the standard scoring function with positive marginally decreasing returns to performing an activity, and constant negative returns to arriving late for an activity and for spending time travelling (Charypar and Nagel 2005).

The other models depart from this base to test the influence of social network scoring, information exchange, and joint evolution of plans and the social network.

- In Config7 agents use standard utility but are free to choose any alternative location for their shopping and leisure. The allocation of the locations in the initial demand with respect to this utility function is probably suboptimal. Agents will redistribute their secondary activities to minimize time losses and to maximize their time spent performing their activities. They will accept tradeoffs in longer trips (time, distance) to avoid congestion.

- Config2 tests time and route planning when friendship with others present is valued as an activity attribute in utility, specifically, the logarithm of the number of friends at a leisure activity. Time coordination of alters' activities may become evident. By offering more utility for leisure activities in which friends are present, a certain insensitivity to time allocations (to the limit that the friends are not missed at the activity) will also be expected. The social network used is random with respect to person attributes or attributes of the social network itself (no edge correlations or other linking preferences), but concentrates friendships into spatially compact areas (see 4.2).

- Config4 alters Config7 to allow secondary locations to be changed, but only for locations learned about through friends. Thus the speed and completeness of knowledge percolation through the social network may limit the learning about space and reduce the activity space relative to the unlimited location choice in Config7, resulting in less satisfactory day plans. One location is exchanged randomly per dyad per iteration. After all agents exchange knowledge, the agents' knowledge is culled to 50% more locations than the agent has need for in its plans. Those locations mentioned most often by friends are the ones most likely to be retained in memory and used for plan optimization. The social network is fixed and identical to that in Config2.

- Config3 combines Config4 and Config2 to allow location choice that is influenced by a fixed social network, plus positive utility for socializing. Thus two feedbacks (replanning and scoring) reinforce the socializing. Larger deviations would be
expected from the time-cost sensitive individual travel behavior defined in the
standard utility function.

- Config5 and Config6 show two versions of co-evolving social networks and travel
  plans to illustrate the ability of the module to create social networks embedded
  geographically in time and space. Friendship will occur in a framework of
  accessibility on the road network, availability of facilities for performing activities,
  and time allocations of other friends. Each model is initialized with the social network
  from Config2. Up to one newly encountered agent per activity per iteration can be
  added as a friend by each ego. The average degree is maintained constant through
  random link removal. All nodes in the social network lose the same proportion of links
  each iteration, on average. Thus high-degree nodes will lose a higher absolute number
  of friends. Establishing and dissolving social connections while learning about space
  through these connections is expected to expand the activity spaces of the agents
  relative to Config4 and Config3, the most similar cases.
4. Results

The simulation was run once for each of the described models, for 500 iterations. The runs were stopped when the executed score (utility) remained constant within 1% over 100 iterations. Ensemble runs with different random seeds would be preferable but for demonstration purposes it is assumed that these results are representative of model behavior under other slightly different random circumstances. MATSim produces output files detailing the micro- and macroscopic spatiotemporal movements of the agents, the evolution of their plans, the state of the traffic flow, the social network topology, and the interactions taking place among befriended agents. Travel simulation results are summarized first, followed by the social network geography.

4.1 Travel simulation results

Table 2 presents the spatiotemporal characteristics of the resulting plans, averaged over all agents in the final iteration, and Figure 1 illustrates the time profiles of the number of agents travelling at any given moment. The social networks enable new ways to spread activities in time and over different routes and to different locations in the agents' attempt to maximize utility. This analysis focuses on linking the different social interaction mechanisms with the observed effect on travel behaviour.
Table 2  Travel behavior summary of seven social network travel models

<table>
<thead>
<tr>
<th>Name</th>
<th>Avg. trip distance (m)</th>
<th>Activity radius (m)</th>
<th>Avg. exec Score</th>
<th>Avg social score</th>
<th>Avg # friends per leisure</th>
<th>(Avg exec score)-(Avg social score)</th>
<th>Avg trip duration (s)</th>
<th>Trip speed (km/h)</th>
<th>Avg distance to friends (m)</th>
<th>Dyad distance (m)</th>
<th>Components</th>
<th>Clust ratio</th>
<th>Diameter</th>
<th>Iterations since last encounter</th>
<th>Meet freq</th>
<th>Avg. link age (iter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Config1</td>
<td>9450</td>
<td>7181</td>
<td>165</td>
<td>0</td>
<td>NA</td>
<td>165</td>
<td>1054</td>
<td>32.3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Config7</td>
<td>11250</td>
<td>8298</td>
<td>166</td>
<td>0</td>
<td>NA</td>
<td>166</td>
<td>1048</td>
<td>38.6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Config2</td>
<td>9470</td>
<td>7181</td>
<td>204</td>
<td>43</td>
<td>0.54</td>
<td>161</td>
<td>1109</td>
<td>30.7</td>
<td>17551</td>
<td>15388</td>
<td>11</td>
<td>3.24</td>
<td>9</td>
<td>500</td>
<td>0.002</td>
<td>500</td>
</tr>
<tr>
<td>Config3</td>
<td>11900</td>
<td>8482</td>
<td>213</td>
<td>55</td>
<td>0.73</td>
<td>158</td>
<td>1274</td>
<td>33.6</td>
<td>17551</td>
<td>15388</td>
<td>11</td>
<td>3.24</td>
<td>9</td>
<td>500</td>
<td>0.002</td>
<td>500</td>
</tr>
<tr>
<td>Config4</td>
<td>11220</td>
<td>8116</td>
<td>165</td>
<td>0</td>
<td>NA</td>
<td>165</td>
<td>1112</td>
<td>36.3</td>
<td>17551</td>
<td>15388</td>
<td>11</td>
<td>3.24</td>
<td>9</td>
<td>500</td>
<td>0.002</td>
<td>500</td>
</tr>
<tr>
<td>Config5</td>
<td>11300</td>
<td>8174</td>
<td>164</td>
<td>0</td>
<td>NA</td>
<td>164</td>
<td>1150</td>
<td>35.4</td>
<td>10162</td>
<td>11396</td>
<td>166</td>
<td>162</td>
<td>12</td>
<td>3.5</td>
<td>0.007</td>
<td>143</td>
</tr>
<tr>
<td>Config6</td>
<td>11320</td>
<td>8165</td>
<td>352</td>
<td>187</td>
<td>5.49</td>
<td>165</td>
<td>1301</td>
<td>31.3</td>
<td>10041</td>
<td>11218</td>
<td>162</td>
<td>160</td>
<td>11</td>
<td>3.4</td>
<td>0.007</td>
<td>143</td>
</tr>
</tbody>
</table>

1 Average distance from agent home location to all visited locations in the plan.
2 Agents retain their 4 best plans each iteration and choose one to execute. Statistics for all plans are available but only the executed plans are comparable with one another.
3 Average distance from agent home location to home locations of all friends (radius of the ego network).
4 Average distance between each pair of friends.
5 Ratio of clustering coefficient to that of an Erdős/Renyi network of the same degree.
Overall, trip distance rises for optimal plans relative to the base case if there is freedom to change secondary locations, regardless of utility function or social interactions. The radius of activities (an indicator of activity spaces) varies less than the trip length and is much smaller than either the world's diameter (100km) or the expected distance between two randomly chosen locations in the world (this is circa 31km if facilities are equally distributed (Dunbar 1997), and in the facilities sample used here it is 36km because of their inhomogenous spatial distribution). Trip travel times are penalized in utility and so depend more on the scoring function than the optimization strategy. Trip speed (an indicator of traffic congestion) is most sensitive to travel time and thus utility. The average utility, net of social interaction, is not directly comparable across scenarios, but it indicates roughly consistent travel times, time delayed, and time participating in the activities.

A static social network with added socializing utility (Config2) but no information exchange yields small changes in routes (increased trip length) and departure time (e.g. more mid-day trips in the time profile) which cause time disutility and lower average speed that is compensated in the social interactions. A static social network with no socializing utility (Config4) but with the ability to change location based on socially mediated information about locations yields similar results to a reference model without social networks and with omniscient knowledge about locations (Config7), namely high speeds, long distances, and fairly efficient coverage of space (i.e. a low ratio of distance travelled to activity radius). Wasted travel time and time not spent at scheduled activities is penalized, so routes and locations are chosen together to spread demand spatially. Adding reinforcement through socializing scoring to this behavior (Config3) does not increase this activity radius very much more, only as much as the travel time can be compensated by increased socializing. The reason the areas and trip distances are consistent between omniscience and socially-mediated knowledge is because the social network has adequate access to the set of optimal locations for the ego with this utility function. Given that this social network enables agents to optimally choose secondary locations, it does not surprise that the travel behavior statistics of the dynamic scenarios Config5 and Config6 also reflect those of the static scenarios Config4 and 3, respectively, with slower speeds and longer travel times emerging when socializing compensates time losses in utility. However, the number of friends encountered per leisure activity in Config6 is dramatically higher than for the other scenarios where this statistic is recorded (section 4.2).
The time profiles of the number of travellers en route shows some compensation effects of the socializing utility score in time, but this varies strongly by configuration. Config1 and Config2 share the inability to adjust locations and both show the most agents starting the day extremely early, probably to compensate for the high demand for travel to the initial locations that are not spread out well, or which have few access routes, concentrating demand onto only a few roads. The socializing reward leads only to slight peak spreading in Config2 relative to Config1. The agents have to be at the same place at the same time in order to benefit from the socializing bonus. This is difficult for them to arrange without making new friends at the new location, or collaborating with existing friends in advance of altering their plans. This effect is probably also behind the profile of Config3 which permits socializing benefits as well as location change, but not new social connections. Without coordinating their moves, the agents have little luck at realizing highly socially rewarding activities that might compensate for suboptimal scheduling. Config6 illustrates how the agents alter their departure times, spreading the travel peaks, when high socializing utility is obtained. They achieve this without coordination through their ability to make new friends (create new utility) at all of their locations, even if they are penalized for extra travel and for missing the ideal (planned) participation time for their activity. Peaks are intensified in Config4, and to a slightly lesser
extent in Config5, which allow changing locations but which do not reward socializing. Here, one sees the demand spreading through space, rather than time, such that activity schedules and low travel times can be maintained (refer to distance in Table 2).

It is not obvious that different agent interactions lie at the base of these differences if one does not know this beforehand.

4.2 Social network and social geography results

Typical social network characteristics and indicators of their geographic embedding are given in Table 2. The statistics are compared to an Erdős/Renyi network of the same number of nodes and average degree. The corresponding expected average dyad distance of a randomly drawn set of pairs of homogenously distributed agents is 31km (see above). The initial spatially embedded social network generated here has an average dyad distance of 15.4km, half as much, because close links are favored. The clustering of the spatially embedded network is 3.24 times higher than that of the non-embedded random network because choosing friends locally means choosing from a smaller pool, where the likelihood is higher to randomly befriend an agent which is already befriended with a friend. The nonspatial random graph would be expected to have a single large component, compared with 11 components of the social network used, here. The expected diameter of the Erdős/Renyi graph is 

\[ (1+O(1))^{*}\left[\frac{\ln(n)}{\ln(k)}\right] \]

where \( n \) is the number of nodes and \( k \) is the graph average degree. For \( n=8760 \) and \( k=12 \), this yields an expected diameter of \( \sim5.5-9 \). The social networks used thus have a slightly longer diameter, and with the higher clustering and more fractured components, are typical *spatially embedded small worlds* with strongly connected local (spatial) neighborhoods and longer connections to more distant neighborhoods (Wong et al. 2006).

The social networks which are permitted to evolve with travel demand and location choice, Config 5 and 6, have much smaller average dyad distances and average distance to friends, longer geodesic distance between neighborhoods, and clustering as well as the number of components that are ten times higher. They reach equilibrium (i.e. the average degree and/or clustering coefficient is constant, Kosinetts 2007 personal communication) after \( \sim80 \) iterations.

Socializing rewards in the utility does not make a difference in the topology of the evolving social networks since Config5 arrived at a graph with similar statistics. The fact that the agents in Config6 met with more friends at once than those in Config5 is still a subject of research but it was not due to the social network topology. It is conceivable that the agents are thinly distributed across leisure facilities in Config5, spread out in space in order to maximize
travel utility, while the socializing utility compensates the agents in Config6 for suboptimal travel plans that let them congregate together at nearly 10 times the density as in Config5.

Figure 3  Final degree distribution for two social network travel models

(a) Static network and initialization, Config2, 3, 4
(b) Network which evolved with activity plan optimization, Config6 (Config5 similar)

Both degree distributions have characteristics typical of small worlds (Poisson-like peak with an exponential tail, Figure 3). The link removal algorithm in Config5 and 6 tends to isolate those agents who do not have many activities or who live in sparsely populated regions. Increased clustering through removal of links on the basis of node degree was also observed by Jin, et. al (2001) in their non-geographic small world simulation. However the degree distribution shifts only slightly to lower degrees and more skewness. The shift in the distance distribution of friendships is much stronger, however, where a spike of small-radius islands forms (Figure 4). The initially exponential dyad distance distribution exhibits a frequency peak at ~4km after evolution.
One question that is interesting to investigate in the models is the investment in travel time and effort needed to maintain a friendship network. The distribution of the average distance from an ego to its friends versus ego degree is very different for the static and dynamic social networks (Figure 5). In Config6 the relationship is roughly a line with positive slope 370m
per friend; i.e. the more friends, the farther they are scattered. In Config2 the distribution is an inverse function of distance. Thus in Config6, the most likely agents to have long-distance connections are those agents with the most connections. This was the opposite in Config2 and in the initialization, where agents with few connections had the longest-distance connections. These could not be economically maintained in the iterative model with link deletion. The long-distance connections, and thus the small-world effects, are not all lost, but the remaining ones are redistributed in a specific way that has to do with geography. Individuals not visited because they live far away, who in reality would be visited less frequently than daily, are (unrealistically) forgotten because of overemphasis on the face to face mechanism of socializing.

Figure 5  Distribution of distance to friends versus number of friends for two social network travel models

(a) Static network and initialization, Config2, 3, 4
(b) Network which evolved with activity plan optimization, Config6 (Config5 similar)
5. Summary and continuing work

The social networks module added to MATSim enables flexible experimentation of travel behavior with social interactions and produces large samples of socializing behavior embedded geographically in time and space. Analysis shows that the manifestations of a "small world" network of social connections become highly varied when embedded within geographic space and time. Simple degree distributions and clustering analyses will not yield understanding of the processes at work. Likewise, it is not always obvious how strongly social interactions determine systemwide travel patterns or are determined by them, even when the interactions are understood at the atomic level in the model specification. The social networks module can help develop intuition and tests for probing data for these phenomena.

The run time and required memory for sizeable scenarios are acceptable for experimentation on small clusters or multi-core computers, even in ensemble configurations. The social network module processes roughly $10^6$ agents per CPU hour and is expected to scale as $\sim 1/(\text{average degree})$ (Marchal 2007).

In the absence of large empirical datasets on socializing and travel behavior, this module is best used to run ensembles with plausible and simple behavioral assumptions to produce distributions of expected outcomes of social and travel interdependence. It would be useful, for example, to simulate a range of social interactions on travel planning and behavior. The systematic study of results would lend insight into recognizing signatures of geography and/or travel behavior in samples of real social networks, the lack of consistent signatures (e.g. behavior subsumed in more easily obtained demographic variables), or the confounding of geographic characteristics with other non-geographical social processes. On the one hand, scalable quantities and universalities may be identified in the system to supplement basic scientific hypotheses in social geography. On the other hand, field work in travel demand could use the guidance to enhance existing travel behavior surveys, e.g. by optimized snowball sampling, or by targeting social clusters, geographical region, transport mode, demographic group, etc. In cases where interactions are suspected, the simulations could provide the intuition for the estimation of correlated utility functions requiring a priori information about social networks.

Other plans include expanding the experimental set to include agent games, such as collaborative activity scheduling. Alternatives to utility maximization are not easily made compatible with MATSim but some approximation of them, such as extremely high rewards or penalties to simulate ultimatum games, may be possible. Implement utility rewards for ride sharing (identical route segments and times) or for meeting together in public transport (with the appropriate multi-modal day plans) would be easy to implement. Realism will be
enhanced in the social network models by including household structures, demographic conditions for befriending (such as homophily indices), and various social processes like incorporating social network structures into the probability of befriending (e.g. triad closure, preferential attachment).

Specific short-term work focuses on the systematic exploration of social network topologies, interactions and parameter values, identifying scaling (universal) effects versus configuration-specific effects, deriving improved measures of the complex phenomena, and experimenting with methods for presentation of the results.
6. References


